心理学报 2022, Vol. 54, No. 9, 1021−1030
Acta Psychologica Sinica

https://doi.org/10.3724/SP.J.1041.2022.01021

声调范畴感知和声调复杂度对音乐音高感知的跨领域影响*

李贤卓[1,2] 肖 容[2] 梁丹丹[2,3]
([1] 南京师范大学国际文化教育学院; [2] 南京师范大学文学院, 南京 210097)
([3] 中国科学技术大学语言交叉研究中心, 合肥 230051)

摘 要 采用音高识别与区分任务, 以汉语母语者、越南语母语者和俄语母语者为研究对象, 考察了两个语言加工层面的因素, 即声调范畴感知模式以及不同语言声调系统的复杂度差异对跨领域音乐音高感知的影响。结果发现: (1)汉语和越南语两组声调语言母语者听辨结果符合范畴感知模式, 而俄语母语者是连续感知。在语言和音乐两种刺激条件下, 声调语言母语者在范畴边界宽度、范畴内区分率、范畴间区分率以及区分峰度等指标上不存在显著差异。(2)汉语和越南语两组声调语言母语者的音乐音高区分任务结果差异不显著。实验结果表明: 在行为层面, 母语声调范畴感知模式可以迁移到音乐音高感知中, 但复杂声调系统并不能促进跨领域的音乐音高精细感知。实验结果从语言对音乐音高加工影响的角度支持了"共享论"。

关键词 范畴感知, 音高, 声调, 音乐
分类号 B842

1 引言

音高是语言和音乐共有的基本声学属性。在语言中, 音高表现为韵律信息, 包括强度、语调以及声调等。声调是一种独特的语言现象, 声调语言通过不同的音高表达不同的词汇语义(Yip, 2002), 如汉语普通话中, 有阴平、阳平、上声和去声四个声调。在音乐中, 音高被记录成不同的音符, 而音高序列构成旋律(Krishnan et al., 2009)。

1.1 语言音高和音乐音高加工的"模块论"与"共享论"

音高在语言和音乐中的表现既有共性又有差异, 对这两个领域的音高加工也存在不同的看法。一些研究支持"模块论"。如 Peretz 和 Coltheart (2003)基于大量脑损伤研究发现, 一个领域存在缺陷的病人另一个领域的功能是完好的。他们认为, 音乐能力与语言能力一样, 有自己的程序和知识库, 由此提出了音乐加工的"模块化模型"。在这个模型中, 领域一般性的基础声学信息经分析后, 由专门的子系统负责处理音乐音高。Chen 等(2016)以及 Chen 等(2018)对声调语言母语者感知音乐音高的研究发现, 这一群体对声调和音乐音高的辨别结果不存在相关性, 他们由此提出"分裂假说(Split Hypothesis)", 认为声调语言母语者声调与非声调音高特征分别表征, 加工时激活不同的神经网络。

相反, 更多的研究支持"共享论"。在语言能力与音乐能力相互联系的基础上, Patel (2008)提出二者可能共享加工资源。基于声调和音乐音高的物理共性, Bradley (2013)认为声调的音高特征(方向、斜率、高度)与音乐的音高特征(轮廓、音程、音调)在声学层面存在映射关系。大量以音乐家为被试的研究认为, 音乐经验和音乐训练有助于提高声调的感知能力(Schön et al., 2004; Marques et al., 2007; Lee & Hung, 2008; Marie et al., 2011)。据此, Besson

收稿日期: 2021-09-16
* 中国博士后科学基金面上项目(2019M651881)和江苏高校哲社项目(2018SJA0211)资助。
通信作者: 梁丹丹, E-mail: ldd233@sina.com

等(2011)提出训练迁移效应，认为语言和音乐共享神经加工机制，长期的音乐经验不但会影响两个领域共同的声学信息加工，也会影响语言的领域特异性知觉。Patel (2012)进一步认为，音乐训练会引发大脑可塑性变化，从而影响语音加工。以上研究立足于从音乐视角看语言加工。Bidelman 等(2013)提出，从理论上看，语言/音乐加工引发的大脑可塑性变化是双向的，声调语言母语者感知音乐音高也具有优势。一些实证研究支持了这一观点，发现无论行为层面还是大脑皮层层面，声调语言母语者对音乐音高的感知优于非声调语言母语者，表明一个领域的优势会迁移至另一领域(Pfordresher & Brown, 2009; Giuliano et al., 2011; Bidelman et al., 2013; Weidema et al., 2016)。上述跨领域研究谈及声调和音乐音高时，更多关注了音高的领域一般性特征，即声学层面上的基频变化，但这些研究并未考虑到声调作为语言要素，在感知上的特异性问题。

1.2 声调范畴感知对音乐音高感知的影响

音位是语言中能够区别意义的最小声音单位，是根据语音的社会属性划分的单位。音位可分为音段音位和超音段音位，前者是从音质角度划分出的不同音位，如语言中不同的元音、辅音，后者是从音质以外的音高、音长等角度划分出的音位，如声调。Liberman 等(1957)发现，英语母语者对于辅音音位的感知属于范畴感知(categorical perception, CP)。范畴感知指听音人受母语音位影响，将音位连续统中多个不同的声音辨认成有限的类别(Liberman et al., 1957)。受该研究启发，后续对超音段音位——声调的研究发现，声调语言母语者对不同声调音位连续统的感知是范畴感知，而非声调语言母语者则是连续感知(Wang, 1976; Hallé et al., 2004; Xu et al., 2006; Peng et al., 2010; Shen & Froud, 2016)。连续感知完全依靠声调的物理属性感知声调，听音人不能将这些等声学间距的音听辨为离散、有限的音位(Xu et al., 2006)。

由此可见，感知声调时，范畴感知模式要求听音人忽视绝大部分刺激间的音高差异，将不同音高的刺激感知为数量有限的音位；但感知音乐音高时，听音人需要区分音高的声学差异，准确地感知每个音符的音高。所以，声调感知与音乐感知存在差异。结合“共享论”，这提示了一个矛盾：如果语言和音乐共享加工机制，声调语言母语者的范畴感知模式可能会迁移到音乐音高感知中，将声学上不同的音高感知为有限的范畴，从而抑制对音乐音高的精细感知。

以往研究从三个角度对这一问题进行过探索，但都存在不足。第一，范畴感知模式的跨领域影响。这类研究发现声调语言母语者对非语言刺激也呈现出类范畴感知模式，暗示声调范畴感知模式的确存在跨领域迁移(Hallé et al., 2004; Xu et al., 2006; Peng et al., 2010)，也就是说，听音人可能会将语言中“不同音高的刺激感知为数量有限的音位”这一模式跨领域迁移到其他领域中，抑制音高精细感知。但这些研究并未明确提出范畴感知模式对音乐音高感知的影响。第二，母语声调的跨领域抑制作用。少数研究注意到母语声调可能存在跨领域抑制作用(Bent et al., 2006; Peretz et al., 2011; Tong et al., 2018)。如 Peretz 等(2011)发现，汉语普通话母语者对小幅度下降音符的感知比非声调语言母语者差，作者认为可能是由于普通话降调通常涉及大范围的基频变化，长期的普通话声调经验降低了被试对降调的敏感度。但是，这些研究并没有明确这种抑制作用与范畴感知的关系。第三，直接考察声调范畴感知模式对音乐音高感知的影响。目前仅发现 Chang 等(2016), Weidema 等(2016)以及 Bidelman 和 Walker (2017)三项研究，但结论相反。前两者发现，声调范畴可能影响音乐音高变化辨认的敏感性，支持存在迁移；后者发现无论是行为层面还是神经层面，汉语母语者对语言感知的范畴化程度都高于音乐，说明范畴感知模式是领域特异性的，不能类推到被试没有经验的领域。值得注意的是，在实验设计上，三项研究都没能很好地匹配两个领域的材料。Chang 等(2016)使用的三音符音乐片段区分任务与典型的范畴感知区分任务差异较大，且音符的音高间隔小，汉语组区分成绩不好很难说是范畴感知的跨领域抑制作用还是实验任务本身难度过大。Weidema 等(2016)对普通话母语者、荷兰语母语非音乐家和荷兰母语音乐家进行的调类识别实验发现，普通话母语者和荷兰语母语的音乐家对上升、下降范畴的感知都好于荷兰语母语非音乐家，但该实验只有识别任务没有区分任务，且其中的“范畴”指领域一般性的类别，如语言、音乐中都存在升、降、平调，无法确定音乐范畴感知是否源于语言迁移。Bidelman 和 Walker (2017)使用的语言材料和音乐材料不匹配。语言材料是从/u/到/a/元音连续统，操纵的变量是第一共振峰而非基频；音乐材料是从大三度到小三度的连续统，操纵的变量是基频。

1.3 母语声调复杂度差异对跨领域音高感知的影响

不同声调语言(方言)的声调复杂度并不一样。汉语普通话有 4 个声调, 一些语言(方言)的声调系统比普通话更为复杂: 泰语有 5 个声调, 粤语、老挝语和越南语都有 6 个声调。这些语言(方言)中存在多个平调、升调或降调, 只是音高不同。一些研究发现声调系统复杂的语言使用者有更精细的音高感知能力(Chen et al., 2012; Zheng et al., 2012; Zheng et al., 2014)。Zheng 等(2014)通过 ERP 研究发现, 相比范畴内刺激, 粤语母语者对非语言刺激(合成锯齿波)的范畴间刺激产生了更强的晚期失匹配成分, 而汉语普通话母语者并没有观察到这一效应, 这暗示粤语母语者将更为精细化的音高感知能力迁移到了非语言感知中, 也就是说声调系统的复杂度可能对被试音高精细感知产生跨领域影响。然而, 我们还未见到相关研究, 能够证实母语声调复杂度对音乐音高感知的影响。

综上所述, 目前对声调语言母语者的跨领域研究虽然揭示了母语经验对音乐音高加工的促进作用, 但多关注领域一般性的音高迁移, 也没有考虑到声调加工中属于语言层面的影响因素, 即范畴感知模式的作用以及不同语言声调系统复杂度的影响。本研究将语音范畴感知连续统应用于音乐音高感知中, 匹配语音刺激与音乐刺激的音高特征, 并以声调范畴相对简单的汉语母语者、声调范畴相对复杂的越南语母语者为研究对象, 同时选取非声调的俄语母语者作为控制组, 讨论以下问题: (1)母语声调范畴感知模式能否迁移到音乐音高感知中? (2)母语声调系统的复杂度差异能否对音乐音高精细感知产生跨领域影响?

对于第一个问题, 本研究预期声调语言母语者感知与母语音高匹配的音乐音高时, 范畴感知模式可以被激活并迁移。对于第二个问题, 我们选取越南语作为复杂声调语言的样本。越南语标准语有 6 个声调调类, 如表 1 所示, 其声学特征, 包括曲折程度、起点终点音高、调值等, 都比汉语普通话更为丰富。本研究预期声调系统的复杂度差异能够跨领域对音乐音高感知产生影响, 越南语母语者对音乐音高的精细感知好于汉语母语者。对上述问题的研究, 可以明确语言音高感知与音乐音高感知的关系, 从声调的语言特异性角度认识“模块论”与“共享论”之争。

表 1　汉语普通话和越南语标准语声调的调类和调值

汉语普通话调类	调值	越南语调类	调值
阴平	55	横声	55
阳平	35	玄声	211
上声	214	问声	213
去声	51	跌声	21-4
		锐声	35
		重声	1

注: 越南语数据源自咸蔓雪(2016)。

2　实验方法

2.1　被试

90 人参与实验, 其中汉语母语者 30 名(以下简称“汉语组”), 14 男 16 女, 平均年龄 21.9 岁, 越南语母语者 30 名(以下简称“越南语组”), 13 男 17 女, 平均年龄 22.1 岁, 俄语母语者 30 名(以下简称“俄语组”), 15 男 15 女, 平均年龄 21.2 岁[1]。越南语组和俄语组都未接受过超过 1 个月的汉语训练。所有被试都未接受过专业机构的声乐或器乐训练, 仅有个别被试接受过业余训练, 且少于 3 年。被试的童年都在所在国度过, 无居住别国大于半年的经历。体检报告表明, 所有被试的听力正常, 没有言语、语言或神经疾病。根据爱丁堡利手量表(Oldfield, 1971), 被试皆为右利手。对 3 组被试的音乐相关能力进行测量: 采用迷你音乐感知技能测试(Mini-PROMS, Zentner & Strauss, 2017)评估音乐能力, 采用蒙特利尔失歌症评估量表(MBEA, Peretz et al., 2003)评估被试是否有失歌症, 未发现音乐能力异常者及失歌症者, 对 3 组被试的评估结果分别进行方差分析, 迷你音乐感知技能测试 $F(2, 89) = 0.62$, $p = 0.540$, 蒙特利尔失歌症评估量表 $F(2, 89) = 1.51$, $p = 0.227$, 差异均不显著。被试参加实验后得到一定的报酬。

2.2　实验材料

实验采用范畴感知范式的识别任务和区分任务。实验材料分为两部分, 一部分为语音刺激, 一部分为音乐刺激。

实验刺激材料所使用的原始语音样本为音节[i¹], 无上下文语境, 由一位普通话二级甲等的男性朗读, 使用 Audition 软件录音, 采样率为 44.1 kHz。使用 Praat 软件, 构造一个由[i¹] (汉语普通话阴平/

[1] 为避免可能存在的方言干扰, 汉语母语者选择北方方言母语者或南方的普通话母语者, 日常使用语言为普通话, 水平为二级甲等及以上。北方方言母语者方言声调不多于越南标准语的 6 个声调。越南语母语者来自河内及附近地区, 日常使用语言为越南标准语。

越南语横声)到[i¹] (汉语普通话阳平/越南语锐声)的连续统: 阴平/横声的起点和终点音高值均设为 225 Hz, 保持终点的音高值不变, 依次降低起点的音高值, 一直到 175 Hz, 跨度范围为 50 Hz, 形成一个包含 9 个刺激的连续统。在合成刺激时, 将刺激起点音高的单位由赫兹(Hz)转换为等效矩形带宽标度(equivalent rectangular bandwidth, ERB), 这是一种心理声学标准, 能真实地反映实际感知到的频率变化情况(Greenwood, 1961)。单个刺激的持续时间为 500 ms, 接近汉语普通话和越南语单音节的时长(郭锦桴, 1993; 潘武俊英, 2008)。使用 Praat 软件中高分辨率的音高提取算法, 分别提取 9 个语音刺激的音高轮廓, 合成同源的正弦波连续音以模拟音乐刺激, 保证模拟的音乐刺激与对应语音刺激的音高轮廓、振幅和持续时间相同。语音刺激和音乐刺激连续统的音高模式如图 1 所示。

表 2 给出两类刺激起始点和终点的音高参数以及两类刺激在各自领域内的实际感知结果。以刺激 9 为例, 该刺激起点音高和终点音高分别为 175 Hz 和 225 Hz, 在语言中被感知为普通话阳平和越南语锐声, 在音乐中近似于一个从 F3 (简谱标记为 4̣) 到 A3 (简谱标记为 6̣)的滑音, 该刺激的波形图和频谱图如图 2 所示。

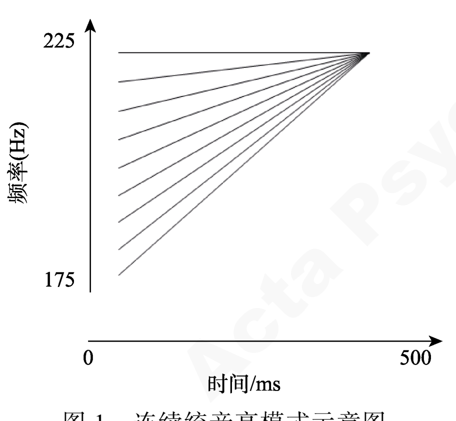


图 1 连续统音高模式示意图

表 2 连续统各刺激的起始音高参数及感知结果

刺激	起始音高(Hz)	起始音高(ERB)	步幅(Hz)	步幅(ERB)	语音刺激感知结果	音乐刺激感知结果
1	225	6.23	6.63	0.12	阴平/横声	A3
2	218.37	6.10	6.52	0.12		A3
3	211.85	5.98	6.41	0.12	↓	G3#→A3
4	205.44	5.86	6.30	0.12		G3#→A3
5	199.14	5.73	6.19	0.12		G3→A3
6	192.95	5.61	6.09	0.12		G3→A3
7	186.87	5.48	5.98	0.12	阳平/锐声	F3#→A3
8	180.88	5.36	5.88	0.12		F3#→A3
9	175	5.24				F3→A3

2.3 实验过程

实验采用经典范畴感知研究的范式, 包括识别任务和区分任务两部分, 在计算机上完成, 材料通过 E-prime 3.0 软件呈现。

识别采用 ABX 任务。被试在每个试次中听到 3 个刺激音, 听到第三个目标刺激音后, 按键判断该刺激的声调与第一个刺激音(语音刺激 1: 阴平/横声; 音乐刺激 1: A3)相同(按“A”键)还是与第二个刺激音(语音刺激 9: 阳平/锐声; 音乐刺激 9: F3→A3)相同(按“L”键)。9 个语音刺激每个出现 10 次, 9 个音乐刺激每个也出现 10 次, 两类刺激共 180 个试次。实验开始时, 在黑屏上显示一个 500 ms 的视觉提示(白色十字), 然后依次播放刺激 1、刺激 9 与目标刺激, 刺激间距为 500 ms。

区分采用 AX 任务。被试在每个试次中听到两个刺激音构成的一个刺激对, 需要判断两个刺激是否相同, 相同按“A”键, 不同按“L”键。刺激对分为两类, 不同刺激对中的两个刺激相差 2 阶, 如连续

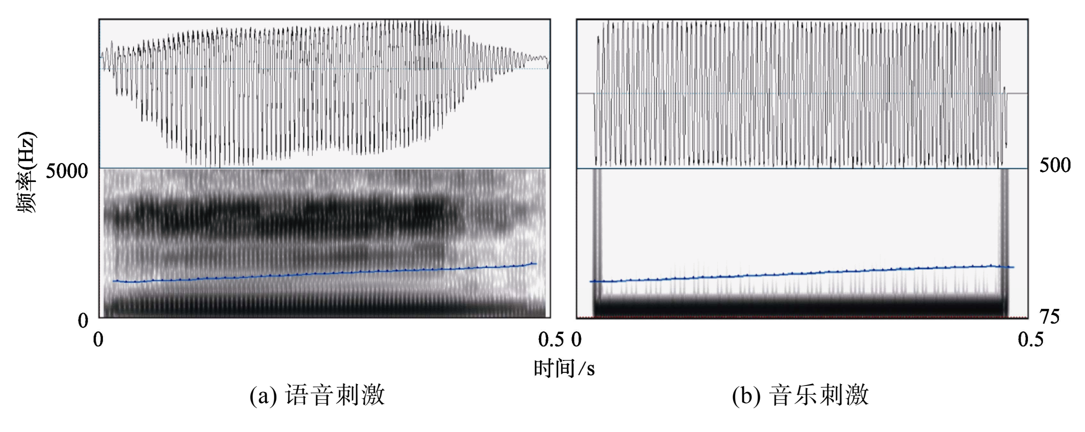


图 2 语音刺激 9 和音乐刺激 9 的波形图(上)和频谱图(下)

统中的第一个和第三个刺激构成刺激对“1-3”, 以此类推, 还有“2-4” “3-5” “4-6” “5-7” “6-8” “7-9”, 每个刺激对重复 10 次, 不同的顺序如“1-3” “3-1”各 5 次; 相同刺激对中的两个刺激完全相同, 从“1-1”到“9-9”。每个刺激对重复 10 次, 刺激间距为 500 ms。语音刺激和音乐刺激各设置 160 个刺激对, 共计 320 个试次。实验开始时, 在黑屏上显示一个 500 ms 的视觉提示(白色十字), 然后播放刺激对。

实验试次以伪随机顺序呈现, 以确保相同试次不连续出现。刺激播放响度为 80 分贝。当被试按键判断或是某个试次呈现 2500 ms 被试还未能做出判断, 屏幕将自动出现 500 ms 的间隔, 然后开始下一试次。随机抽取一半被试先进行语言任务, 再进行音乐任务; 另一半被试先进行音乐任务, 再进行语言任务。为使被试熟悉实验, 在每部分正式实验之前先练习, 练习采用的刺激与正式实验不同。被试可以一直进行练习, 直到完全掌握实验要求。

2.4　数据统计

根据 Xu 等(2006)、Peng 等(2010)、Jiang 等(2012)的方法, 识别任务的评价标准为识别回归曲线的斜率和范畴边界宽度两个指标, 区分任务的评价标准为范畴内区分率、范畴间区分率和峰度三个指标。

将识别任务结果进行 logistic 回归分析, 得到回归方程: $\ln\frac{P_1}{1-P_1}=b_0+b_1x$。其中 P_1 是被试将实验刺激判定为阴平/横声/A3 的概率, x 是 9 个刺激, b_0 是回归曲线的截距, b_1 是回归曲线的斜率。b_1 绝对值越大, 范畴化程度越高。识别曲线的范畴边界宽度为 $P_1=0.25$ 和 $P_1=0.75$ 之间的线性距离, 范畴边界宽度越小, 范畴化程度越高。

区分任务中, 将所有刺激对分为 7 个刺激对组, 每个刺激对组包含四类刺激对。如, 刺激对组“1-3”包含“1-3” “3-1” “1-1” “3-3”四类刺激对。数据以刺激对组为单位进行统计。计算各刺激对组的区分率 $P_{(i,j)}$。计算公式为: $P_{(i,j)}=P\frac{\text{"S"}}{S}\bullet P(S)+P\frac{\text{"D"}}{D}\bullet P(D)$。其中, S 是刺激对组中, 两个刺激相同的个数, D 是两个刺激不同的个数, P (S)是相同刺激对的个数与刺激对总数之比, P (D)是不同刺激对的个数与刺激对总数之比, $P\frac{\text{"S"}}{S}$ 是被试做出“相同”判断的刺激对与所有相同的刺激对数量之比, $P\frac{\text{"D"}}{D}$ 是被试做出“不同”判断的刺激对与所有不同的刺激对数量之比。按照上述方法, 计算每个刺激对组的正确率。范畴间区分率为跨越识别曲线范畴边界位置(识别率为 0.5 时, 对应的识别曲线横坐标值)区分对组的区分率均值, 如被试识别曲线范畴边界为 5.1, 那么该被试范畴间区分率为跨越该位置的区分对组“4-6”和“5-7”的平均区分正确率。范畴内区分率为该被试其他 5 个刺激对组的区分正确率均值。峰度是范畴间区分率与范畴内区分率之差。范畴化程度越高, 范畴内区分率越低, 而范畴间区分率和峰度越高。

3　实验结果

3.1　识别任务

三组被试语言和音乐刺激的识别率图像如图 3 所示, 汉语组和越南语组被试语音刺激和音乐刺激的识别率较为接近, 在连续统起始部分和终点部分的几个刺激分别接近 100%和 0, 中间部分出现突变, 从 100%骤降至接近 0, 基本符合范畴感知识别任务的特征, 而俄语组无论是语音刺激还是音乐刺激, 都与两个声调母语组差距较大, 接近于连续感知。

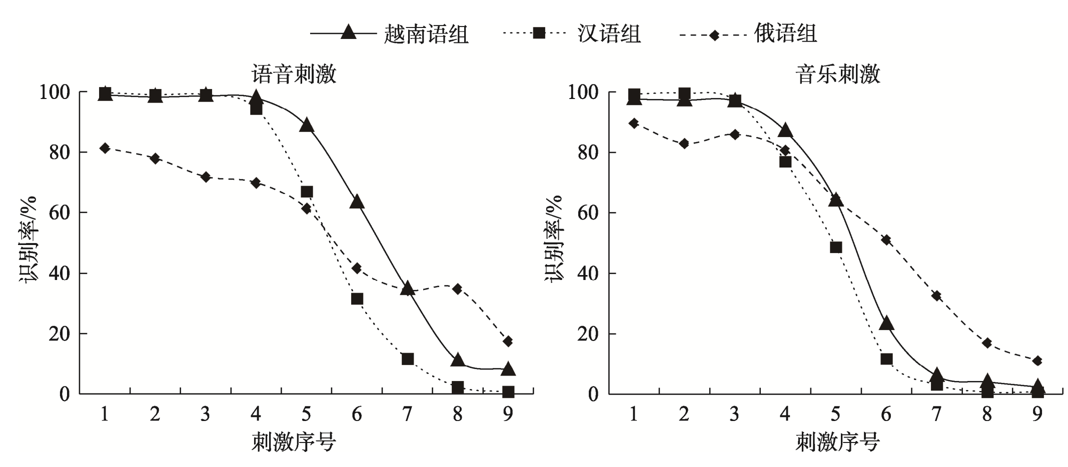


图 3　三组被试识别任务结果示意图

三组被试识别曲线的斜率、范畴边界宽度如表3所示。

以组别(越南语、汉语、俄语)作为组间因素，刺激对类型(语音、音乐)作为组内因素，对识别曲线斜率、范畴边界宽度分别进行3×2的重复测量方差分析。结果显示，在识别曲线的斜率上，刺激对类型的主效应显著，$F(1, 87) = 6.19$, $p = 0.015$, $\eta_p^2 = 0.07$。事后分析显示，语音刺激识别曲线的斜率显著小于音乐刺激，$p = 0.015$, 95% CI = [–0.03, 0.27]。组别主效应显著，$F(2, 87) = 58.09$, $p < 0.001$, $\eta_p^2 = 0.57$。刺激对类型和组别的交互效应显著，$F(2, 87)=3.29$, $p = 0.042$。组别在刺激对类型上的简单效应分析表明，在语音刺激上，汉语组显著好于越南语组，$p = 0.002$, 95% CI = [0.22, 0.90]；汉语组显著好于俄语组，$p < 0.001$, 95% CI = [–1.39, –0.71]；越南语组显著好于俄语组，$p < 0.001$, 95% CI = [–1.94, –1.26]。在音乐刺激上，汉语组与越南语组差异不显著，$p = 0.186$；汉语组显著好于俄语组，$p < 0.001$, 95% CI = [–1.72, –1.03]；越南语组显著好于俄语组，$p < 0.001$, 95% CI = [–1.95, –1.26]。在边界宽度上，刺激对类型的主效应不显著，$F(1, 87) = 0.87$, $p = 0.352$。组别主效应显著，$F(2, 87) = 3.40$, $p = 0.038$, $\eta_p^2 = 0.07$。事后分析显示，汉语组与越南语组差异不显著，$p = 0.980$；汉语组显著高于俄语组，$p = 0.027$, 95% CI = [–33.01, –2.21]；越南语组显著高于俄语组，$p = 0.026$, 95% CI = [–32.82, –2.01]。刺激对类型和组别的交互效应不显著，$F(2, 87) = 0.83$, $p = 0.352$。这些结果表明，汉语组和越南语组更为接近，他们与俄语组差异较大。

表 3 三组被试识别任务相关指标(n = 90)

组别	语音刺激		音乐刺激	
	斜率	范畴边界宽度	斜率	范畴边界宽度
越南语组	–1.83(0.45)	1.28(0.34)	–2.20(0.58)	1.06(0.27)
汉语组	–2.39(0.70)	1.02(0.38)	–2.43(0.54)	0.94(0.17)
俄语组	–0.79(0.79)	26.69(99.29)	–0.83(0.84)	10.48(14.92)

注：括号内是标准差。

3.2 区分任务

三组被试语言和音乐刺激对的区分结果如图4所示，区分任务各指标如表4所示。

以组别(越南语、俄语、汉语)作为组间因素，刺激对类型(语音、音乐)作为组内因素，对范畴内区分率、范畴间区分率、峰度分别进行3×2的重复测量方差分析。结果显示，在范畴内区分率上，刺激对类型的主效应不显著，$F(1, 87) = 0.001$, $p = 0.972$。组别主效应显著，$F(2, 87) = 4.95$, $p = 0.009$, $\eta_p^2 = 0.10$。刺激对类型和组别的交互效应不显著，$F(2, 87) = 0.92$, $p = 0.404$。事后分析显示，越南语组

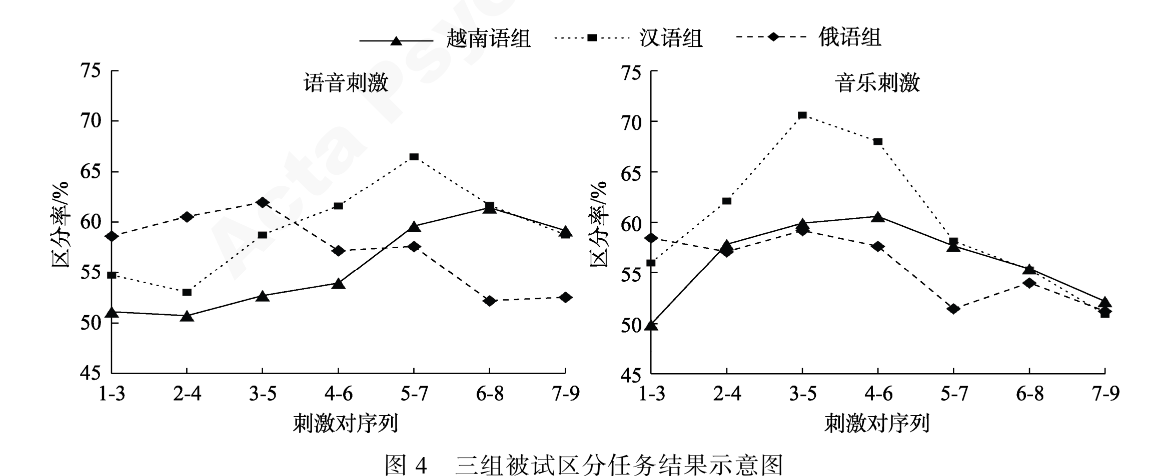


图 4 三组被试区分任务结果示意图

表 4 三组被试区分任务相关指标(n = 90)

组别	语音刺激			音乐刺激		
	范畴内区分率	范畴间区分率	峰度	范畴内区分率	范畴间区分率	峰度
越南语组	0.55 (0.07)	0.60 (0.09)	0.05 (0.07)	0.56 (0.05)	0.56 (0.09)	0.07 (0.09)
汉语组	0.55 (0.12)	0.60 (0.15)	0.03 (0.17)	0.58 (0.06)	0.66 (0.11)	0.12 (0.12)
俄语组	0.51 (0.13)	0.49 (0.24)	–0.07 (0.25)	0.48 (0.21)	0.51 (0.25)	–0.04 (0.23)

注：括号内是标准差。

显著高于俄语组, p = 0.017, 95% CI = [0.01, 0.11]; 汉语组显著高于俄语组, p = 0.004, 95% CI = [0.02, 0.12]; 汉语组与越南语组差异不显著, p = 0.615。在范畴间区分率上, 刺激对类型的主效应不显著, F(1, 87) = 0.36, p = 0.550。组别主效应显著, F(2, 87) = 8.27, p = 0.001, η_p^2 = 0.16。刺激对类型和组别的交互效应不显著, F(2, 87) = 0.87, p = 0.422。事后分析显示, 越南语组显著高于俄语组, p = 0.009, 95% CI = [0.02, 0.15]; 汉语组显著高于俄语组, p < 0.001, 95% CI = [0.07, 0.20]; 汉语组与越南语组差异不显著, p = 0.184。刺激对类型和组别的交互效应不显著, F(2, 87) = 0.87, p = 0.422。在峰度上, 刺激对类型的主效应不显著, F(1, 87) = 3.52, p = 0.064。组别主效应显著, F(2, 87) = 8.80, p < 0.001, η_p^2 = 0.17。事后分析显示, 越南语组显著高于俄语组, p = 0.002, 95% CI = [0.04, 0.17]; 汉语组显著高于俄语组, p < 0.001, 95% CI = [0.07, 0.20]; 汉语组与越南语组差异不显著, p = 0.487。刺激对类型和组别的交互效应不显著, F(2, 87) = 0.87, p = 0.422。以上结果表明, 两个声调母语组区分任务的结果基本接近, 与俄语组差异较大。

4　讨论

本研究考察了声调感知中的语言特异性因素对音乐音高感知的跨领域影响, 选取母语声调类型相对简单、声调类型相对复杂以及非声调语言母语者作为被试, 讨论语言加工层面范畴感知模式以及声调系统复杂度的跨领域迁移。

4.1　声调范畴感知模式的跨领域迁移

第一个问题讨论母语声调范畴感知模式能否迁移到音乐音高感知中。实验结果发现, 无论在语言还是音乐条件下, 汉语、越南语两个声调母语组的识别曲线都出现了突变, 区分曲线都出现了区分峰, 都呈现出了范畴感知模式。汉语组和越南语组在范畴边界宽度、范畴内区分率、范畴间区分率以及区分峰度等指标上都不存在语言与音乐的刺激类型差异, 说明母语范畴感知模式可以迁移到音乐音高感知中, 这一结论与预期一致。而俄语组识别曲线未能出现明显突变, 区分曲线较为平缓并出现多个区分峰, 在五项指标上与两个声调母语组差异显著, 在语言音乐两类刺激上不存在显著差异, 说明俄语组依靠刺激的声学属性, 对语音和音乐刺激做出了较为一致的连续感知。

以往从语言音高感知到音乐音高感知的迁移研究, 实验任务多为两个音符或音乐片段的音高差异区分(Pfordresher & Brown, 2009; Giuliano et al., 2011; Hutka et al., 2015), 结果发现声调语言母语者的表现好于非声调语言母语者, 这是由于声调母语音高曲拱的加工经验提高了对音高感知的敏感性, 被试在完成实验任务时只需要关注音高的物理差异这一因素(Giuliano et al., 2011)。与上述研究不同, 本研究主要考察属于音位层面的范畴感知模式这一声调加工因素的迁移, 发现在音乐音高的识别和区分任务中, 两个声调母语组都呈现出了与语音刺激类似的模式, 说明声调范畴感知模式可以迁移到音乐音高感知中。本研究的结果支持了“共享论”。与本研究结果不同, 前人研究中, 有结果表明声调语言母语者在音乐音高感知中并不存在优势。如 Pfordresher 和 Brown (2009)的研究发现, 虽然声调语言母语者总体上对音乐音高的感知好于非声调语言母语者, 但在音高区分任务中的表现与非声调语言母语者差异不显著, 作者因此认为音乐和语言共享资源, 但不共享表征(Patel, 2008)。本研究认为, 得出此结果的原因可能就在于上述研究仅关注了音高的物理差异, 而未考虑到语言层面因素的影响。在语言层面范畴感知模式下, 声调语言母语者虽能够很好地感知范畴间刺激的差异, 但不能区分范畴内刺激, 将不同的音乐刺激感知为相同的范畴内刺激, 抑制了音高的精细感知。Pfordresher 等人的研究未考虑到这种可能性, 故得出与本研究不同的结论。尽管也有学者提出了“母语影响”“抑制”(Bent et al., 2006; Tong et al., 2018)等概念, 但并未进行详细讨论。具体来说, 这种“母语影响”“抑制”可能源于母语声调范畴的跨领域迁移。就本实验而言, 如果孤立地看区分任务中两个声调组的范畴内区分率与俄语组差异并不显著这一结果, 极有可能忽略范畴感知模式的跨领域抑制作用, 错误地认为声调语言母语者的音高加工优势未能迁移到音乐音高感知中, 得出支持“模块论”的结论。总之, 进行声调母语者的音乐音高感知结果解释时, 应充分考虑到范畴感知因素的作用。

本研究的结果与 Chang 等(2016)、Weidema 等(2016)基本一致而与 Bidelman 和 Walker (2017)不同。这可能是由于 Bidelman 等人的研究并未做到语音刺激与音乐刺激的匹配。一般情况下, 语言中的声调与音乐中的音符分属不同的表征, 声调以范畴化的方式表征, 而音乐音级以离散的方式表征(Bradley, 2013), 音乐家不能将音程知识直接迁移

到声调感知中，声调语言母语者也不能将范畴感知模式直接迁移到音乐旋律感知中。只有一个领域获得的特定音高范畴经验与另一个领域需要加工的音高信息相匹配，这种促进效应才可能发生(Asaridou & McQueen, 2013)。Bidelman 和 Walker (2017)使用的语言材料是元音连续统，音乐材料是从大三度到小三度的连续统，二者并不同质，音高信息也不能匹配，被试在完成音乐音高区分任务时不能激活语音的范畴感知模式。

4.2 母语声调系统的复杂度差异对音乐音高感知的影响

本研究的第二个问题是母语声调系统的复杂度差异能否影响音乐音高精细度感知。对音高感知精细度问题的考察一般采用音高区分任务(Pfordresher & Brown, 2009; Giuliano et al., 2011; Hutka et al., 2015)。本实验中的区分任务测量被试对不同音高的语音/乐音的区分程度，可以反映被试对音高精细度差异的感知。但研究结果发现，越南语组与汉语组无论是范畴内还是范畴间区分率，在语言和音乐音高条件下差异都不显著，表明母语的复杂声调系统并不能促进跨领域的音乐音高精细感知。实验结果未能支持预期。

导致这一结果的原因可能有二。第一，从语言角度看，本研究的语言材料是汉语普通话阴平−阳平/越南语横声−锐声连续统，越南语组被试在感知时激活的越南语声调表征与汉语组被试激活的普通话声调表征基本一致，因此两组被试范畴间区分率不存在显著差异。较之汉语普通话，越南语声调的复杂性体现在玄声、跌声等调类上，本研究的实验材料并不包含这些声调，故而未能检验出越南语被试复杂声调系统带来的更为精细的声调加工能力。Peng 等(2010)和 Zheng 等(2012)对比了普通话和声调系统更为复杂的粤语，也得出了类似结论。他们的行为研究发现，感知汉语普通话音高时，粤语母语者与普通话母语者在声学层面未见差异。但是，Zheng 等(2012)的 ERP 实验发现，在语言内部，粤语母语者的精细音高感知优势可以向汉语普通话声调感知迁移，这种影响仅发生在大脑皮质层面而不能体现在行为层面。未来可开展神经电生理研究，进一步探究母语声调复杂度差异对音乐音高感知的跨领域影响问题。第二，从音乐角度看，本研究结果发现，两组声调语言母语者音乐刺激的范畴间区分率都显著高于俄语母语者，这表明声调语言经验对于跨领域的音乐音高区分有显著的增益作用，从语言到音乐这一方向支持了“共享论”以及 Besson 等(2011)的训练迁移效应说。但根据 OPERA 假说(Patel, 2012)，音乐经验比语言经验更能够促进精细的音高感知，因为音乐感知需要不断调节和监控音调，对编码精度的要求更高。以往实证研究也发现类似的结论，在辨别一些以声学信息为基础的非母语声调上，声调语言母语者不如音乐家(Cooper & Wang, 2012; Chang et al., 2016)。因此，尽管越南语的声调系统较汉语普通话更为复杂，但与音乐旋律相比，越南语调类数量和音高精细程度仍显简单，因此越南语母语者感知音乐音高时，在行为层面没能产生跨领域的促进作用。

5 结论

本研究考察了影响声调加工的语言学因素对音乐音高感知的跨领域迁移。研究发现，声调的范畴感知模式可以迁移到音乐音高感知中；复杂的母语声调并不能促进跨领域音乐音高精细感知。研究结果从语言对音乐加工影响的角度支持了“共享论”。

参 考 文 献

Asaridou, S. S., & Mcqueen, J. M. (2013). Speech and music shape the listening brain: Evidence for shared domain-general mechanisms. *Frontiers in Psychology, 4*(4), 1–14.

Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology Human Perception & Performance, 32*(1), 97–103.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology, 2*, 1–11.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLoS ONE, 8*(4), Article e60676.

Bidelman, G. M., & Walker, B. S. (2017). Attentional modulation and domain-specificity underlying the neural organization of auditory categorical perception. *European Journal of Neuroscience, 45*(5), 690–699.

Bradley, E. D. (2013). Pitch perception in lexical tone and melody. *Reviews of Research in Human Learning and Music, 1*: Epub2013002.

Chang, D., Hedberg, N., & Wang, Y. (2016). Effects of musical and linguistic experience on categorization of lexical and melodic tones. *Journal of the Acoustical Society of America, 139*(5), 2432–2447.

Chen, A., Liu, L., & Kager, R. (2016). Cross-domain correlation in pitch perception, the influence of native language. *Language, Cognition and Neuroscience, 31*(6), 751–760.

Chen, A., Peter, V., Wijnen, F., Schnack, H., & Burnham, D. (2018). Are lexical tones musical? Native language's influence on neural response to pitch in different domains. *Brain and Language, 180–182*(15), 31–41.

Chen, Z., Peng, L., Wang, E. Q., Larson, C. R., Huang, D., & Liu, H. (2012). ERP correlates of language-specific processing of auditory pitch feedback during self-vocalization. *Brain & Language, 121*(1), 25–34.

Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *Journal of the Acoustical Society of America, 131*(6), 4756–4769.

Giuliano, R. J., Pfordresher, P., Stanley, E. M., Narayana, S., & Wicha, N. Y. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in Psychology, 2*, 1–12.

Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America, 33*(10), 1344–1356.

Guo, J. F. (1993). *Probe into Chinese tone and intonation*, Beijing: Beijing Language Institute Press.

[郭锦桴. (1993). *汉语声调语调阐要与探索*. 北京：北京语言学院出版社.]

Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*(3), 395–421.

Hutka, S., Bidelman, G. M., & Moreno, S. (2015). Pitch expertise is not created equal: Cross-domain effects of musicianship and tone language experience on neural and behavioural discrimination of speech and music. *Neuropsychologia, 71*, 52–63.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Memory & Cognition, 40*(7), 1109–1121.

Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience, 21*(6), 1092–1105.

Lee, C.-Y., & Hung, T.-H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *Journal of the Acoustical Society of America, 124*(5), 3235–3248.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*(5), 358–368.

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience, 23*(10), 2701–2715.

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19*(9), 1453–1463.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.

Patel, A. D. (2012). The opera hypothesis: Assumptions and clarifications. *Annals of the New York Academy of Sciences, 1252*(1), 124–128.

Peng, G., Zheng, H.-Y., Gong, T., Yang, R.-X., Kong, J.-P., & Wang, W. S.-Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics, 38*(4), 616–624.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders——The montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences, 999*, 58−75.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688–691.

Peretz, I., Nguyen, S., & Cummings, S. (2011). Tone language fluency impairs pitch discrimination. *Frontiers in Psychology, 2*, 1–5.

Pfordresher, P., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics, 71*(6), 1385–1398.

Phan, V. T. (2008). An analysis of the acoustic features of static tones in Vietnamese. *Journal of Yunnan Normal University (Teaching Chinese as a Foreign Language and Research Edition)*, *6*(4), 59–65.

[潘武俊英. (2008). 越南语静态声调的声学分析. *云南师范大学学报(对外汉语教学与研究版)*, *6*(4), 59-65.]

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341–349.

Shen, G., & Froud, K. (2016). Categorical perception of lexical tones by English learners of Mandarin Chinese. *Journal of the Acoustical Society of America, 140*(6), 4396–4403.

Tong, X., Choi, W., & Man, Y. Y. (2018). Tone language experience modulates the effect of long-term musical training on musical pitch perception. *Journal of the Acoustical Society of America, 144*(2), 690–697.

Wang, W. S. (1976). Language change. *Annals of the New York Academy of Sciences, 280*(1), 61–72.

Weidema, J. L., Roncaglia-Denissen, M. P., & Honing, H. (2016). Top-down modulation on the perception and categorization of identical pitch contours in speech and music. *Frontiers in Psychology, 7*, 1–11.

Xian, M. X. (2016). *Historical hierarchical analysis of relational morphemes in Chinese and Vietnamese*. Shanghai: Zhongxi Publishing House.

[咸蔓雪. (2016). *汉语越南语关系语素历史层次分析*. 上海：中西书局.]

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America, 120*(2), 1063–1074.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the short‐proms and the mini‐proms. *Annals of the New York Academy of Sciences, 1400*(1), 33–45.

Zheng, H.-Y., Minett, J. W., Peng, G., & Wang, W. S.-Y. (2012). The impact of tone systems on the categorical perception of lexical tones: An Event-Related Potentials study. *Language and Cognitive Processes, 27*(2), 184–209.

Zheng, H.-Y., Peng, G., Chen, J., Zhang, C., Minett, J. W., & Wang, W. S. (2014). The influence of tone inventory on ERP without focal attention: A cross-language study. *Computational and Mathematical Methods in Medicine*, Article 961563.

The cross-domain influence of tonal categorical perception and tonal complexity on musical pitch perception

LI Xianzhuo[1,2], XIAO Rong[2], LIANG Dandan[2,3]

([1] *International College for Chinese Studies, Nanjing Normal University, Nanjing* 210097, *China*)

([2] *School of Chinese Language and Literature, Nanjing Normal University, Nanjing* 210097, *China*)

([3] *Interdisciplinary Research Center for Linguistic Science, University of Science and Technology of China, Hefei* 230051, *China*)

Abstract

Pitch is a fundamental acoustic property shared by both language and music. However, there are different views on the processing of tonal pitch and musical pitch. Some studies support the "modularity view", suggesting that tonal pitch and musical pitch are represented separately. In contrast, more studies support the "shared domain-general view", implying that there are commonalities in the processing of tonal pitch and musical pitch based on their physical similarities. The existing studies have mostly focused on domain-general pitch transfer; nonetheless, they have not considered domain specificity of the tonal pitch as a linguistic element. Hence, it remains unclear whether domain specificity factors, such as categorical perception and complexity of different tonal language experiences on speakers' pitch processing, play a role in musical pitch perception.

To address the above issues, ninety participants were involved in the experiment, including native speakers of Mandarin Chinese (with a relatively simple pitch category), native speakers of Vietnamese (with a relatively complex pitch category), and native speakers of Russian (nonpitched control group). A 3 (group: Vietnamese vs. Chinese vs. Russian) × 2 (stimulus type: speech vs. music) between-and-within-subjects design was used. A continuum from [i˥] (closed for Yinping in Chinese and Transverse in Vietnamese) to [i˧˥] (closed for Yangping in Chinese and Acute in Vietnamese) and its musical counterpart was constructed. Participants were first tested in an ABX identification task to determine whether the stimulus X was similar to A or B and then an AX discrimination task to decide whether the two stimuli in the pair were identical. The experiment was performed in E-prime 3.0.

The results showed that (1) both the Chinese and Vietnamese groups showed categorical perception for language and music pitches. There were no differences between language and music stimuli in category boundary width, within-category discrimination rate, between-category discrimination rate, or discrimination peak for either group. The Russian group's identification curve did not show abrupt shifts, and their discrimination curve was relatively flat with multiple peaks, which indicates a continuous pattern significantly different from the two tonal language groups. (2) There were no significant differences between the Vietnamese group and the Chinese group in either the within-category discrimination rate or the between-category discrimination rate.

The experimental results suggest that at the behavioral level, the pattern of native tonal categorical perception can transfer to musical pitch perception, but tonal complexity does not facilitate cross-domain musical pitch perception. The findings of this study support the "shared domain-general view" in terms of the influence of language on musical pitch processing.

Key words categorical perception, pitch, tone, music